经 典 原 版 书 库

离散时间控制系统

（英文版·第2版）

Second Edition

Discrete-Time Control Systems

Katsuhiko Ogata

Katsuhiko Ogata 著

明尼苏达大学

离散时间控制系统

（英文版·第2版）

Discrete-Time Control Systems
(Second Edition)

Katsuhiko Ogata
明尼苏达大学　著

Preface

This book presents a comprehensive treatment of the analysis and design of discrete-time control systems. It is written as a textbook for courses on discrete-time control systems or digital control systems for senior and first-year graduate level engineering students.

In this second edition, some of the older material has been deleted and new material has been added throughout the book. The most significant feature of this edition is a greatly expanded treatment of the pole-placement design with minimum-order observer by means of the state-space approach (Chapter 6) and the polynomial-equations approach (Chapter 7).

In this book all materials are presented in such a way that the reader can follow the discussions easily. All materials necessary for understanding the subject matter presented (such as proofs of theorems and steps for deriving important equations for pole placement and observer design) are included to ease understanding of the subject matter presented.

The theoretical background materials for designing control systems are discussed in detail. Once the theoretical aspects are understood, the reader can use MATLAB with advantage to obtain numerical solutions that involve various types of vector-matrix operations. It is assumed that the reader is familiar with the material presented in my book *Solving Control Engineering Problems with MATLAB* (Prentice Hall) or its equivalent.

The prerequisites for the reader are a course on introductory control systems, a course on ordinary differential equations, and familiarity with MATLAB computations. (If the reader is not familiar with MATLAB, it may be studied concurrently.)

Since this book is written from the engineer's point of view, the basic concepts involved are emphasized and highly mathematical arguments are carefully avoided in the presentation. The entire text has been organized toward a gradual development of discrete-time control theory.

The text is organized into eight chapters and three appendixes. The outline of the book is as follows: Chapter 1 gives an introduction to discrete-time control systems. Chapter 2 presents the z transform theory necessary for the study of discrete-time control systems. Chapter 3 discusses the z plane analysis of discrete-time systems, including impulse sampling, data hold, sampling theorem, pulse transfer function, and digital filters. Chapter 4 treats the design of discrete-time control systems by conventional methods. This chapter includes stability analysis of closed-loop systems in the z plane, transient and steady-state response analyses, and design based on the root-locus method, frequency-response method, and analytical method.

Chapter 5 presents state-space analysis, including state-space representations of discrete-time systems, pulse transfer function matrix, discretization method, and Liapunov stability analysis. Chapter 6 discusses pole-placement and observer design. This chapter contains discussions on controllability, observability, pole placement, state observers, and servo systems. Chapter 7 treats the polynomial equations approach to control systems design. This chapter first discusses the Diophantine equation and then presents the polynomial equations approach to control systems design. Finally, model matching control systems are designed using the polynomial equations approach. Chapter 8 presents quadratic optimal control. Both finite-stage and infinite-stage quadratic optimal control problems are discussed. This chapter concludes with a design problem based on quadratic optimal control solved with MATLAB.

Appendix A presents a summary of vector-matrix analysis. Appendix B gives useful theorems of the z transform theory that were not presented in Chapter 2, the inversion integral method, and the modified z transform method. Appendix C discusses the pole-placement design problem when the control signal is a vector quantity.

Examples are presented at strategic points throughout the book so that the reader will have a better understanding of the subject matter discussed. In addition, a number of solved problems (A problems) are provided at the end of each chapter, except Chapter 1. These problems represent an integral part of the text. It is suggested that the reader study all these problems carefully to obtain a deeper understanding of the topics discussed. In addition, many unsolved problems (B problems) are provided for use as homework or quiz problems.

Most of the materials presented in this book have been class-tested in senior and first-year graduate level courses on control systems at the University of Minnesota.

All the materials in this book may be covered in two quarters. In a semester course, the instructor will have some flexibility in choosing the subjects to be covered. In a quarter course, a good part of the first six chapters may be covered. An instructor using this text can obtain a complete solutions manual from the

publisher. This book can also serve as a self-study book for practicing engineers who wish to study discrete-time control theory by themselves.

Appreciation is due to my former students who solved all the solved problems (A problems) and unsolved problems (B problems) and made numerous constructive comments about the material in this book.

Katsuhiko Ogata

Contents

Chapter 3

z-Plane Analysis of Discrete-Time Control Systems   74

Chapter 4

Design of Discrete-Time Control Systems by Conventional Methods   173

Chapter 5

State-Space Analysis   293

Chapter 6

Pole Placement and Observer Design   377

Chapter 7

Polynomial Equations Approach to Control Systems Design   517

Chapter 8

Quadratic Optimal Control Systems   566

Appendix A

Vector-Matrix Analysis   633

Appendix B

z Transform Theory   681

Appendix C

Pole Placement Design with Vector Control   704

1

Introduction to Discrete-Time Control Systems

1–1 INTRODUCTION

In recent years there has been a rapid increase in the use of digital controllers in control systems. Digital controls are used for achieving optimal performance—for example, in the form of maximum productivity, maximum profit, minimum cost, or minimum energy use.

Most recently, the application of computer control has made possible "intelligent" motion in industrial robots, the optimization of fuel economy in automobiles, and refinements in the operation of household appliances and machines such as microwave ovens and sewing machines, among others. Decision-making capability and flexibility in the control program are major advantages of digital control systems.

The current trend toward digital rather than analog control of dynamic systems is mainly due to the availability of low-cost digital computers and the advantages found in working with digital signals rather than continuous-time signals.

Types of Signals. A continuous-time signal is a signal defined over a continuous range of time. The amplitude may assume a continuous range of values or may assume only a finite number of distinct values. The process of representing a variable by a set of distinct values is called *quantization*, and the resulting distinct values are called *quantized* values. The quantized variable changes only by a set of distinct steps.

An analog signal is a signal defined over a continuous range of time whose amplitude can assume a continuous range of values. Figure 1–1(a) shows a continuous-time analog signal, and Figure 1–1(b) shows a continuous-time quantized signal (quantized in amplitude only).

1
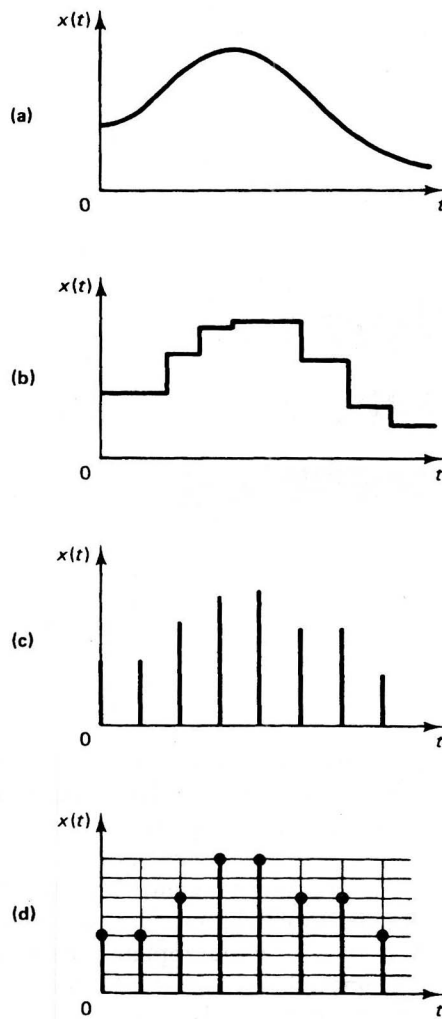
Figure 1-1 (a) Continuous-time analog signal; (b) continuous-time quantized signal; (c) sampled-data signal; (d) digital signal.

Notice that the analog signal is a special case of the continuous-time signal. In practice, however, we frequently use the terminology "continuous-time" in lieu of "analog." Thus in the literature, including this book, the terms "continuous-time signal" and "analog signal" are frequently interchanged, although strictly speaking they are not quite synonymous.

A discrete-time signal is a signal defined only at discrete instants of time (that is, one in which the independent variable t is quantized). In a discrete-time signal, if the amplitude can assume a continuous range of values, then the signal is called a *sampled-data signal*. A sampled-data signal can be generated by sampling an analog signal at discrete instants of time. It is an amplitude-modulated pulse signal. Figure 1-1(c) shows a sampled-data signal.

A digital signal is a discrete-time signal with quantized amplitude. Such a signal can be represented by a sequence of numbers, for example, in the form of binary

numbers. (In practice, many digital signals are obtained by sampling analog signals and then quantizing them; it is the quantization that allows these analog signals to be read as finite binary words.) Figure 1–1(d) depicts a digital signal. Clearly, it is a signal quantized both in amplitude and in time. The use of the digital controller requires quantization of signals both in amplitude and in time.

The term "discrete-time signal" is broader than the term "digital signal" or the term "sampled-data signal." In fact, a discrete-time signal can refer either to a digital signal or to a sampled-data signal. In practical usage, the terms "discrete time" and "digital" are often interchanged. However, the term "discrete time" is frequently used in theoretical study, while the term "digital" is used in connection with hardware or software realizations.

In control engineering, the controlled object is a plant or process. It may be a physical plant or process or a nonphysical process such as an economic process. Most plants and processes involve continuous-time signals; therefore, if digital controllers are involved in the control systems, signal conversions (analog to digital and digital to analog) become necessary. Standard techniques are available for such signal conversions; we shall discuss them in Section 1–4.

Loosely speaking, terminologies such as discrete-time control systems, sampled-data control systems, and digital control systems imply the same type or very similar types of control systems. Precisely speaking, there are, of course, differences in these systems. For example, in a sampled-data control system both continuous-time and discrete-time signals exist in the system; the discrete-time signals are amplitude-modulated pulse signals. Digital control systems may include both continuous-time and discrete-time signals; here, the latter are in a numerically coded form. Both sampled-data control systems and digital control systems are discrete-time control systems.

Many industrial control systems include continuous-time signals, sampled-data signals, and digital signals. Therefore, in this book we use the term "discrete-time control systems" to describe the control systems that include some forms of sampled-data signals (amplitude-modulated pulse signals) and/or digital signals (signals in numerically coded form).

Systems Dealt With in This Book.   The discrete-time control systems considered in this book are mostly linear and time invariant, although nonlinear and/or time-varying systems are occasionally included in discussions. A linear system is one in which the principle of superposition applies. Thus, if y_1 is the response of the system to input x_1 and y_2 the response to input x_2, then the system is linear if and only if, for every scalar α and β, the response to input $\alpha x_1 + \beta x_2$ is $\alpha y_1 + \beta y_2$.

A linear system may be described by linear differential or linear difference equations. A time-invariant linear system is one in which the coefficients in the differential equation or difference equation do not vary with time, that is, one in which the properties of the system do not change with time.

Discrete-Time Control Systems and Continuous-Time Control Systems.
Discrete-time control systems are control systems in which one or more variables can change only at discrete instants of time. These instants, which we shall denote by kT or t_k ($k = 0, 1, 2, \ldots$), may specify the times at which some physical measurement

is performed or the times at which the memory of a digital computer is read out. The time interval between two discrete instants is taken to be sufficiently short that the data for the time between them can be approximated by simple interpolation.

Discrete-time control systems differ from continuous-time control systems in that signals for a discrete-time control system are in sampled-data form or in digital form. If a digital computer is involved in a control system as a digital controller, any sampled data must be converted into digital data.

Continuous-time systems, whose signals are continuous in time, may be described by differential equations. Discrete-time systems, which involve sampled-data signals or digital signals and possibly continuous-time signals as well, may be described by difference equations after the appropriate discretization of continuous-time signals.

Sampling Processes.   The sampling of a continuous-time signal replaces the original continuous-time signal by a sequence of values at discrete time points. A sampling process is used whenever a control system involves a digital controller, since a sampling operation and quantization are necessary to enter data into such a controller. Also, a sampling process occurs whenever measurements necessary for control are obtained in an intermittent fashion. For example, in a radar tracking system, as the radar antenna rotates, information about azimuth and elevation is obtained once for each revolution of the antenna. Thus, the scanning operation of the radar produces sampled data. In another example, a sampling process is needed whenever a large-scale controller or computer is time-shared by several plants in order to save cost. Then a control signal is sent out to each plant only periodically and thus the signal becomes a sampled-data signal.

The sampling process is usually followed by a quantization process. In the quantization process the sampled analog amplitude is replaced by a digital amplitude (represented by a binary number). Then the digital signal is processed by the computer. The output of the computer is sampled and fed to a hold circuit. The output of the hold circuit is a continuous-time signal and is fed to the actuator. We shall present details of such signal-processing methods in the digital controller in Section 1–4.

The term "discretization," rather than "sampling," is frequently used in the analysis of multiple-input–multiple-output systems, although both mean basically the same thing.

It is important to note that occasionally the sampling operation or discretization is entirely fictitious and has been introduced only to simplify the analysis of control systems that actually contain only continuous-time signals. In fact, we often use a suitable discrete-time model for a continuous-time system. An example is a digital-computer simulation of a continuous-time system. Such a digital-computer-simulated system can be analyzed to yield parameters that will optimize a given performance index.

Most of the material presented in this book deals with control systems that can be modeled as linear time-invariant discrete-time systems. It is important to mention that many digital control systems are based on continuous-time design techniques. Since a wealth of experience has been accumulated in the design of continuous-time

controllers, a thorough knowledge of them is highly valuable in designing discrete-time control systems.

1-2 DIGITAL CONTROL SYSTEMS

Figure 1–2 depicts a block diagram of a digital control system showing a configuration of the basic control scheme. The system includes the feedback control and the feedforward control. In designing such a control system, it should be noted that the "goodness" of the control system depends on individual circumstances. We need to choose an appropriate performance index for a given case and design a controller so as to optimize the chosen performance index.

Signal Forms in a Digital Control System.   Figure 1–3 shows a block diagram of a digital control system. The basic elements of the system are shown by the blocks. The controller operation is controlled by the clock. In such a digital control system, some points of the system pass signals of varying amplitude in either continuous time or discrete time, while other points pass signals in numerical code, as depicted in the figure.

The output of the plant is a continuous-time signal. The error signal is converted into digital form by the sample-and-hold circuit and the analog-to-digital converter. The conversion is done at the sampling time. The digital computer
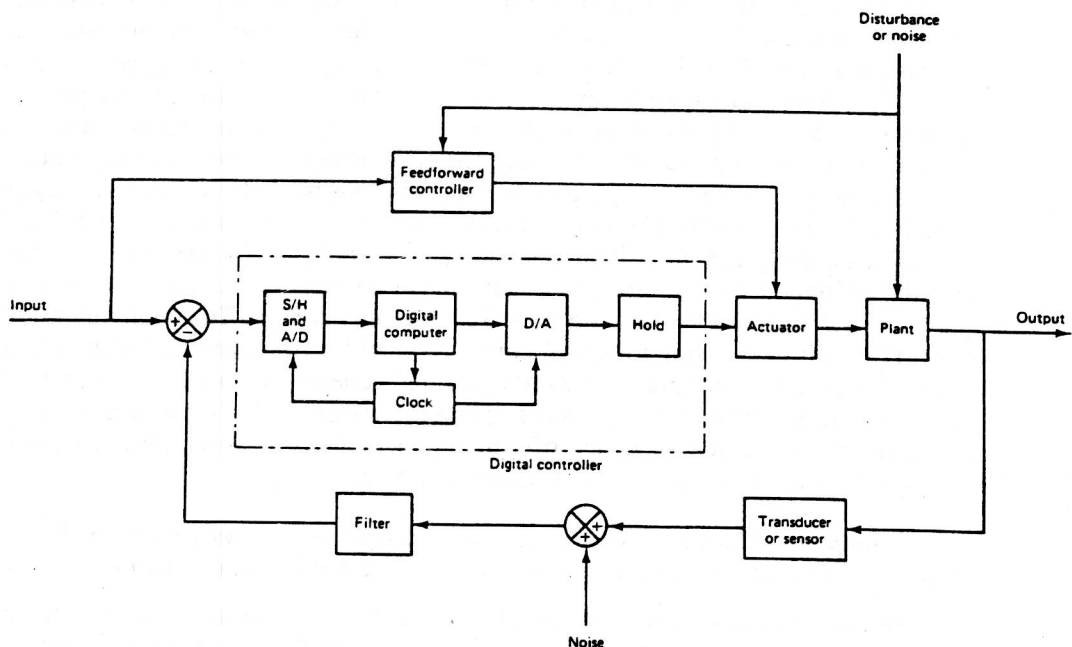

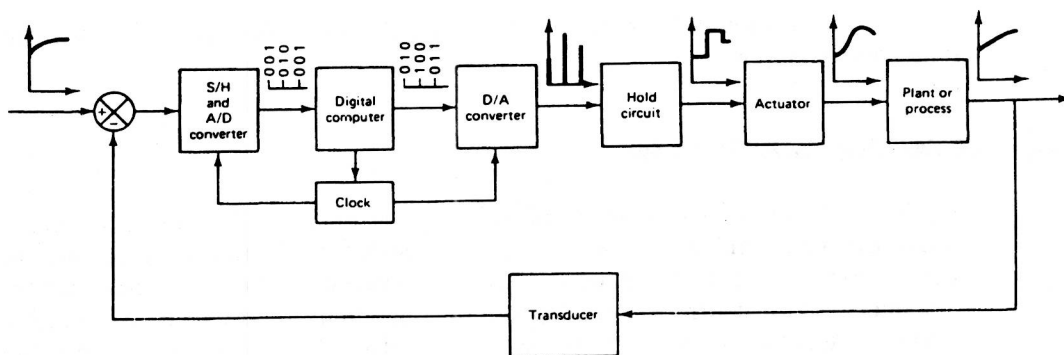**Figure 1–2** Block diagram of a digital control system.

Figure 1–3 Block diagram of a digital control system showing signals in binary or graphic form.

processes the sequences of numbers by means of an algorithm and produces new sequences of numbers. At every sampling instant a coded number (usually a binary number consisting of eight or more binary digits) must be converted to a physical control signal, which is usually a continuous-time or analog signal. The digital-to-analog converter and the hold circuit convert the sequence of numbers in numerical code into a piecewise continuous-time signal. The real-time clock in the computer synchronizes the events. The output of the hold circuit, a continuous-time signal, is fed to the plant, either directly or through the actuator, to control its dynamics.

    The operation that transforms continuous-time signals into discrete-time data is called *sampling* or *discretization*. The reverse operation, the operation that transforms discrete-time data into a continuous-time signal, is called *data-hold*; it amounts to a reconstruction of a continuous-time signal from the sequence of discrete-time data. It is usually done using one of the many extrapolation techniques. In many cases it is done by keeping the signal constant between the successive sampling instants. (We shall discuss such extrapolation techniques in Section 1–4.)

    The sample-and-hold (S/H) circuit and analog-to-digital (A/D) converter convert the continuous-time signal into a sequence of numerically coded binary words. Such an A/D conversion process is called *coding* or *encoding*. The combination of the S/H circuit and analog-to-digital converter may be visualized as a switch that closes instantaneously at every time interval T and generates a sequence of numbers in numerical code. The digital computer operates on such numbers in numerical code and generates a desired sequence of numbers in numerical code. The digital-to-analog (D/A) conversion process is called *decoding*.

    Definitions of Terms.   Before we discuss digital control systems in detail, we need to define some of the terms that appear in the block diagram of Figure 1–3.

    Sample-and-Hold (S/H).   "Sample-and-hold" is a general term used for a sample-and-hold amplifier. It describes a circuit that receives an analog input signal and holds this signal at a constant value for a specified period of time. Usually the signal is electrical, but other forms are possible, such as optical and mechanical.

Analog-to-Digital Converter (A/D).   An analog-to-digital converter, also called an encoder, is a device that converts an analog signal into a digital signal, usually a numerically coded signal. Such a converter is needed as an interface between an analog component and a digital component. A sample-and-hold circuit is often an integral part of a commercially available A/D converter. The conversion of an analog signal into the corresponding digital signal (binary number) is an approximation, because the analog signal can take on an infinite number of values, whereas the variety of different numbers that can be formed by a finite set of digits is limited. This approximation process is called *quantization*. (More on quantization is presented in Section 1–3.)

Digital-to-Analog Converter (D/A).   A digital-to-analog converter, also called a decoder, is a device that converts a digital signal (numerically coded data) into an analog signal. Such a converter is needed as an interface between a digital component and an analog component.

Plant or Process.   A plant is any physical object to be controlled. Examples are a furnace, a chemical reactor, and a set of machine parts functioning together to perform a particular operation, such as a servo system or a spacecraft.

A process is generally defined as a progressive operation or development marked by a series of gradual changes that succeed one another in a relatively fixed way and lead toward a particular result or end. In this book we call any operation to be controlled a process. Examples are chemical, economic, and biological processes.

The most difficult part in the design of control systems may lie in the accurate modeling of a physical plant or process. There are many approaches to the plant or process model, but, even so, a difficulty may exist, mainly because of the absence of precise process dynamics and the presence of poorly defined random parameters in many physical plants or processes. Thus, in designing a digital controller, it is necessary to recognize the fact that the mathematical model of a plant or process in many cases is only an approximation of the physical one. Exceptions are found in the modeling of electromechanical systems and hydraulic-mechanical systems, since these may be modeled accurately. For example, the modeling of a robot arm system may be accomplished with great accuracy.

Transducer.   A transducer is a device that converts an input signal into an output signal of another form, such as a device that converts a pressure signal into a voltage output. The output signal, in general, depends on the past history of the input.

Transducers may be classified as analog transducers, sampled-data transducers, or digital transducers. An analog transducer is a transducer in which the input and output signals are continuous functions of time. The magnitudes of these signals may be any values within the physical limitations of the system. A sampled-data transducer is one in which the input and output signals occur only at discrete instants of time (usually periodic), but the magnitudes of the signals, as in the case of the analog transducer, are unquantized. A digital transducer is one in which the input and output signals occur only at discrete instants of time and the signal magnitudes are quantized (that is, they can assume only certain discrete levels).